



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION PAPERS

10

OF

PAUL KIMELMAN

15

AND

IAN FIELD

20

FOR

25

ACCESS TO BIT VALUES WITHIN DATA WORDS

STORED IN A MEMORY

BACKGROUND OF THE INVENTION

Field of the Invention

5 This invention relates to the field of data processing systems. More particularly, this invention relates to access to bit values within data words stored in a memory.

Description of the Prior Art

10 It is known to provide data processing systems with memories for storing data values to be manipulated. Such memories are typically arranged to store data values as 8-bit byte values. These 8-bit byte values (or greater) are written together and read together.

15 Within many data processing systems, there is a requirement to store data values that are smaller than one byte in length. Using a full byte of storage within a memory to store only one or a few bits of data is disadvantageously inefficient in terms of circuit area, cost, power consumption and in other ways. Furthermore, many non-memory devices require access to specific bits within control and status registers.

20 In order to reduce memory storage requirements it is known to store multiple data values within a single byte of memory storage. However, a disadvantage of this approach is that when it is desired to write or read such a data value from within a byte, additional processing steps are required to extract the portion of the data value
25 concerned or to insert the portion of the data value concerned into the byte being accessed within the memory. These extra instructions disadvantageously slow execution and increase code size.

30 One known technique for addressing this problem is to add special purpose instructions for inserting certain bits or extracting certain bits into a byte or word value stored within a memory. However, the instruction bits space consumed by these special purpose instructions is a disadvantage. Furthermore, these are usually instructions used in addition to load and store so extra processing time is required. These extra instructions, whether separate instructions or part of the load and store,

can still create read-modify-write problems due to interrupts splitting the read from the modify and then write.

Another known approach is that of the SHARC processor in which the same
5 underlying memory can be accessed as either 16-bit, 32-bit or 40-bit words.

A further known approach is the Intel 8051 which provides a bit memory, but lacks the flexibility of a normal memory system.

10 SUMMARY OF THE INVENTION

Viewed from one aspect the present invention provides an apparatus for processing data, said apparatus comprising:

a memory operable to store data values; and

15 memory accessing logic responsive to memory access instructions to access data values stored within said memory; wherein

said memory has a first memory address region and a second memory address region;

said memory accessing logic is operable, in response to a memory access
20 instruction specifying a first memory access to a first data value within said first memory address region, to convert said first memory access to a second memory access, said second memory access being to a second data value within said second memory address region;

when said first memory access is a memory write, said second memory access
25 is a read-modify-write memory access in which Y bits within said first data value are written to Y bits within said second data value with those bits within said second data value other than said Y bits being unaltered; and

when said first memory access is a memory read, said second memory access is a masked read memory access in which Y bits of said first data value are read from
30 Y bits of said second data value and those bits within said first data value other than said Y bits are set to a predetermined value independent of bits of said second data value other than said Y bits.

The present technique provides a way in which standard memory access instructions can be used to address selected bit values within a stored data value through the use of a special purpose memory address region in which differing memory locations map to different bit portions of data values stored in another memory address region. This technique can be considered to obviate the need for special purpose bit accessing instructions by allowing the memory address being accessed within the special purpose memory region to effectively encode the additional bit addressing information within a standard memory access instruction and so allow this to be mapped to the desired bit accessing operation within the data values where the bits concerned are actually stored. The memory access to the bit values which is generated is in itself not a standard memory access in that for a write operation it performs a read-modify-write access as is appropriate to avoid altering other data stored within the memory location concerned and for a read access it performs a masked read so as to pick out only the bit values desired.

It will be appreciated that the number of bits it is desired to access within a data word can vary and in preferred embodiments this may be a user programmable value in the range between 1 and 8.

The data values addressed by the memory accesses can vary in size but will typically be 32, 16 or 8-bits in length.

Whilst it is possible that the bits to be accessed within a data value could be spread in arbitrary positions throughout that data value, it is simpler when the bits to be accessed together are contiguous within the data value concerned. In this way, the bottom 3, 4 or 5-bits of the first address encodes the starting bit in all cases.

The processing of data values to be written or read from bits within a stored data value is simplified when those bit values form the least significant bits of the first data value.

The present technique has particular applicability when the memory is a byte addressable memory and the number of bits to be accessed in one operation is less than eight.

In order to simplify interrupt behaviour, preferred embodiments are such that the read-modify-write memory access is performed as an atomic operation.

5 It will be appreciated that since the first memory access is converted to another form prior to being physically issued on an associated memory bus structure, it is possible that the first memory access is unaligned with the second memory access being appropriately realigned to match the bus structure provided.

10 Viewed from another aspect the present invention provides a method of processing data, said method comprising the steps of:

storing data values within a memory; and

in response to memory access instructions, accessing data values stored within said memory; wherein

15 said memory has a first memory address region and a second memory address region;

in response to a memory access instruction specifying a first memory access to a first data value within said first memory address region, converting said first memory access to a second memory access, said second memory access being to a second data value within said second memory address region;

20 when said first memory access is a memory write, said second memory access is a read-modify-write memory access in which Y bits within said first data value are written to Y bits within said second data value with those bits within said second data value other than said Y bits being unaltered; and

25 when said first memory access is a memory read, said second memory access is a masked read memory access in which Y bits of said first data value are read from Y bits of said second data value and those bits within said first data value other than said Y bits are set to a predetermined value independent of bits of said second data value other than said Y bits.

30 The above, and other objects, features and advantages of this invention will be apparent from the following detailed description of illustrative embodiments which is to be read in connection with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 schematically illustrates a data processing system including a memory for storing data values;

5

Figure 2 schematically illustrates a portion of the memory address space/map for the system of Figure 1;

Figure 3 schematically illustrates the mapping between a first memory address region and a second memory address region and;

10

Figure 4 schematically illustrates the relationship between a first data value and a second data value;

15

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Figure 1 schematically illustrates a data processing system 2 including a processor core 4, a data memory 6 and an instruction memory 8.

The processor core 4 includes processing logic including a register bank 10, a multiplier 12, a shifter 14 and an adder 16. The processing logic 10, 12, 14, 16 operates under control of control signals generated by an instruction decoder 18 which is itself responsive to instructions within an instruction pipeline 20. The instructions loaded to the pipeline 20 are retrieved from the instruction memory 8.

25

The instruction decoder 18 is responsive to memory access instructions to generate control signals which operate to control a load store unit 22 within the processor core 4. The load store unit 22 performs data value read and data write operations within the data memory 6. In accordance with the techniques described below, the load store unit 22, or alternatively a separate unit between the load store unit 22 and the data memory 6, provides a conversion of memory accesses made to a first memory address region into memory accesses to a second memory address region that pick out particular bits within the data values stored in the data memory 6.

30

Figure 2 schematically illustrates a portion of the memory address map for the system of Figure 1. This is the data address map and includes a base data region 24, an extra data region 26 and a bit-band data region 28. The load store unit 22 is operable to interpret/convert memory access instructions to the bit-band data region 28 into memory accesses to the base data region 24. Each byte within the bit-band data region 28 in this example maps to a particular bit within a byte value in the base data region 24. In the example processor of Figure 1, the data accesses are 8-bit data accesses with the least significant bit of a data access to the bit-band data region 28 mapping to a particular bit within a byte data value of the base data region 24. Thus, by addressing an appropriate byte within the bit-band data region 28, a programmer can effectively access an individual bit within a data value within the base data region 24.

Figure 3 schematically illustrates this relationship between data accesses to the bit-band data region 28 and the base data region 24 in more detail. As illustrated, a first data access to the bit-band data region 28 is an 8-bit data access. The least significant bit Y of this data access is the bit which is to be mapped to a bit within the base data region 24. This mapping is performed by the load store unit circuitry 22 of Figure 1 which performs the conversion illustrated mathematically at the bottom of Figure 3 to map between byte addresses within the bit-band data region 28 and the byte and bit within to be accessed within the base data region 24.

In different embodiments in which the accesses to the bit band data region are 32-bit word accesses, as part of this operation, the accesses may be converted from a 32-bit accesses to 8-bit access, although this change in access size is optional and in some embodiments the data accesses may more conveniently be left as a 32-bit data accesses to match the existing bus structures and for other reasons. The data accesses may also be realigned as required to match any constraints of the bus structures. As an example, a 32-bit unaligned access may be mapped to a 32-bit aligned access.

In the case of a write being made to the bit-band data region 28, this is mapped into an atomic read-modify-write operation to the base data region 24. The atomic nature of this read-modify-write operation means that it is not possible to interrupt this operation part way through and this is important in order to maintain data integrity and consistency when dealing with interrupt processing. A read operation to the bit-band

data region 28 maps to a masked read operation from the base data region 24 such that the appropriate bit within the byte of the base data region 24 is picked out and returned as the least significant bit within the 8-bit result which appears to the program to have been retrieved from the bit-band data region 28. The bit picked out appears in the least significant bit position of the returned data value, irrespective of its position within the byte of the base data region 24. The bit to be written in the case of write accesses also appears at the least significant bit position within the data word written to the bit-band data region 28 even though it maps to a arbitrary position within the byte of the base data region 24.

Figure 4 schematically illustrates the relationship between the data value accessed within the bit-band data region 28 and the data value accessed within the base data region 24 in an alternative way. It will be seen that the data values in these two regions can have differing and effectively arbitrary lengths N, M chosen to suit the particular requirements or to match the existing requirements of the system concerned. The bit or bits being accessed are mapped as illustrated and the relationship can effectively be viewed as a funnel with the bits other than the Y bits to be accessed being set to predetermined values or left unaltered as appropriate depending upon whether the operation is a read or a write. It will be appreciated that in the case of a read of a particular bit value, the returned data value can have its bit positions other than the selected bits being addressed set to a predetermined value, such as 0. Alternatively the bit positions may be used to sign extend the returned data value.

Although illustrative embodiments of the invention have been described in detail herein with reference to the accompanying drawings, it is to be understood that the invention is not limited to those precise embodiments, and that various changes and modifications can be effected therein by one skilled in the art without departing from the scope and spirit of the invention as defined by the appended claims.